

JOHANNES FEHRENBACH and LUDGER RÜSCHENDORF (Freiburg)

**ANALYSIS OF MARKOV CHAIN ALGORITHMS  
ON SPANNING TREES, ROOTED FORESTS,  
AND CONNECTED SUBGRAPHS**

*Abstract.* We analyse a natural edge exchange Markov chain on the set of spanning trees of an undirected graph by the method of multicommodity flows. The analysis is then refined to obtain a canonical path analysis. The construction of the flow and of the canonical paths is based on related path constructions in a paper of Cordovil and Moreira (1993) on block matroids. The estimates of the congestion measure imply a polynomial bound on the mixing time. The canonical paths for spanning trees also yield polynomial time mixing rates for some related Markov chains on the set of forests with roots and on the set of connected spanning subgraphs. We obtain a parametric class of stationary distributions from which we can efficiently sample. For rooted forests this includes the uniform distribution. For connected spanning subgraphs the uniform distribution is not covered.

**1. Introduction.** Counting the number of spanning trees in an undirected, connected, cycle-free graph is one of the few counting problems on graphs  $G = (V, E)$  which can be solved deterministically in polynomial time. Recall that a *spanning tree* is a maximum cardinality cycle-free subgraph  $(V, S)$  of  $G$ . If  $D = \text{diag}(d_0, \dots, d_{n-1})$  is the diagonal matrix with the degrees  $d_0, \dots, d_{n-1}$  of the  $n$  vertices in  $V$  on its main diagonal and if  $A$  denotes the adjacency matrix of  $V$ , then the classical Kirchhoff formula states that for all  $0 \leq i \leq n - 1$ ,

$$(1.1) \quad \# \text{ spanning trees of } G = \det (D - A)_{ii}$$

where  $(D - A)_{ii}$  is the  $(n - 1) \times (n - 1)$  principal submatrix of  $D - A$  obtained by deleting the  $i$ th row and the  $i$ th column. Since the determinant

---

2000 *Mathematics Subject Classification*: 68W20, 68W40.

*Key words and phrases*: spanning trees, randomized algorithm, multicommodity flow, canonical paths, Markov chain, rooted forests, connected subgraphs.

of a matrix may be computed in time  $O(n^3)$  by the Gaussian elimination algorithm this formula implies a polynomial time algorithm for counting spanning trees in an undirected graph. Various techniques have been developed to establish randomized approximation schemes (RAS) in particular for the approximate counting of combinatorial and graph-theoretic structures. For a survey see Jerrum and Sinclair (1996) or Jerrum (2003). These *Markov chain Monte Carlo methods* have been successfully applied to establish approximate polynomial time counting algorithms for a series of difficult counting and approximation problems like the number of perfect matchings, the graph colouring problem, the approximation of the partition function in the Ising model, the volume of convex bodies, and many others.

The basic problem is as follows. Let  $\Omega = \Omega_n$  be a finite set depending on the length  $n$  of the *input* (like the number of nodes in a graph-theoretic problem). Let  $N = N_n$  be the (unknown) number of elements of  $\Omega$ . Then a random algorithm  $A = A_n$  is called a *fully polynomial randomized approximation scheme* (FPRAS) for  $N$  if for all  $n$  and for all small  $\varepsilon, \delta \in (0, 1)$ ,

$$(1.2) \quad P((1 - \varepsilon)N \leq A \leq (1 + \varepsilon)N) \geq 1 - \delta$$

where the algorithm runs in time bounded by a polynomial in  $n, \varepsilon^{-1}$  and  $\delta^{-1}$ . Note that  $N = N(n, \varepsilon, \delta)$  and  $A = A(n, \varepsilon, \delta)$  depend on  $n, \varepsilon, \delta$ , and typically  $N$  grows exponentially fast in  $n$ . Therefore, a naive Monte Carlo algorithm will not work in this case. The dependence on  $\delta$  can be easily neglected by repeated sampling.

Jerrum, Valiant, Vazirani suggested a RAS by reducing the problem of approximate counting of  $N = |\Omega|$  to that of almost uniform random sampling in the following way:

Assume that there is some decreasing sequence of subsets  $\Omega = \Omega_0 \supset \Omega_1 \supset \dots \supset \Omega_r$  with the following properties:

- (a)  $|\Omega_r|$  can be calculated,
- (b)  $|\Omega_i|/|\Omega_{i+1}|$  is polynomially bounded in  $n$  for  $0 \leq i < r$ ,
- (1.3) (c)  $r = r_n$  is polynomially bounded in  $n$ ,
- (d) for  $0 \leq i < r$  elements of  $\Omega_i$  can be sampled approximately uniformly in polynomial time with respect to  $n$ .

Then by iterative sampling in  $\Omega_i$  as in (d) one obtains an estimator  $\bar{X}_i$  of  $|\Omega_{i+1}|/|\Omega_i|$  for  $0 \leq i < r$ . By (b) a polynomial number of samples in  $n$  and  $\varepsilon^{-1}$  is sufficient. Define the RAS  $A$  by

$$(1.4) \quad A := |\Omega_r| |\bar{X}|^{-1}$$

where  $\bar{X} := \prod_{i=0}^{r-1} \bar{X}_i$ . By (a) and (c),  $\bar{X}$  can be calculated in polynomial time. Typically it will not be difficult to construct natural subsequences  $\Omega_i$  such that (a)–(c) hold. In typical applications the  $\Omega_i$  are of the same structure as  $\Omega$  (with different size). So solving (d) for  $\Omega_0 = \Omega$  yields solutions

of (d) for any  $\Omega_i$ . This property is called *self-reducibility* in Jerrum *et al.* (1986). But the challenge is condition (d) for  $\Omega = \Omega_0$ . The main tool for constructing approximately uniform samples is to construct a suitable Markov chain  $\mathfrak{M}$  which has as its stationary distribution the uniform distribution on  $\Omega$ . (A well known application of this method is the simulated annealing algorithm.)

The main technical problem is to establish that the Markov chain used is rapidly mixing. This property implies that sampling can be done efficiently (in polynomial time). Several tools have been developed for this purpose. The main ones are various eigenvalue estimates for the second largest eigenvalue of the transition matrix (Diaconis and Stroock (1991)), coupling methods (see e.g., Aldous (1983) or Diaconis (1988)), the conductance method and the method of canonical paths (see Diaconis and Stroock (1991) and Sinclair (1993)). Here the basic idea to obtain a good upper bound on the mixing time of the Markov chain is to select a set  $\Gamma = \{\gamma_{xy} \mid x, y \in \Omega\}$  of *canonical paths*  $\gamma_{xy}$  for each pair  $(x, y)$  such that no transition  $e = (v, w)$  of the graph of the chain is used too often in the corresponding flow problem. Define the *congestion measure*  $\varrho = \varrho(\Gamma)$  by

$$(1.5) \quad \varrho(\Gamma) = \max_{e=(v,w) \in E} \frac{1}{\pi(v)P(v,w)} \sum_{(x,y) : \gamma_{xy} \ni e} \pi(x)\pi(y)|\gamma_{xy}|.$$

Here  $P(v, w)$  is the transition matrix of the Markov chain,  $\pi$  is the stationary distribution and  $|\gamma_{xy}|$  is the length of the path  $\gamma_{xy}$ . The sum term is the total flow through edge  $e$  while the first term is the inverse capacity of edge  $e$ . By a result of Diaconis and Stroock (1991) one obtains, for irreducible, aperiodic, reversible Markov chains with  $P(x, x) \geq 1/2$  for all  $x \in \Omega$ , the following estimate of the mixing time  $\tau = \tau(\varepsilon)$ :

$$(1.6) \quad \tau(\varepsilon) \leq \varrho(\Gamma)(\log \hat{\pi}^{-1} + \log \varepsilon^{-1})$$

with  $\hat{\pi} = \min_{x \in \Omega} \pi(x)$ . So the problem of bounding the mixing time of the Markov chain can be reduced to obtaining bounds on the congestion measure. Note that for the uniform distribution  $\pi$  the term  $\log \hat{\pi}^{-1}$  is polynomial in  $n$ , while typically  $N = |\Omega|$  is exponential in  $n$ .

To obtain bounds on  $\varrho(\Gamma)$  an important technique is to *code* the transitions of the canonical path. For  $e \in T = \{(x, y) \in \Omega \times \Omega \mid P(x, y) > 0\}$  define  $\mathcal{P}_c(e) = \{(x, y) \in \Omega \times \Omega \mid \gamma_{xy} \ni e\}$ , the set of pairs whose canonical path uses edge  $e$ . Then a *coding* is a system  $(\eta_e)$  of injective mappings  $\eta_e : \mathcal{P}_c(e) \rightarrow \Omega$ ,  $e \in T$ . If  $\pi$  is the uniform distribution on  $\Omega$  and  $(\eta_e)$  a coding, then we obtain the estimate

$$(1.7) \quad \varrho(\Gamma) = \max_{e=(v,w) \in T} \frac{1}{\pi(v)P(v,w)} \sum_{(x,y) \in \mathcal{P}_c(e)} \pi(x)\pi(y)|\gamma_{xy}|$$

$$\begin{aligned} &\leq \max_{e=(v,w) \in T} \frac{\ell_{\max}}{P(v,w)} \sum_{(x,y) \in \mathcal{P}_c(e)} \pi(\eta_e(x,y)) \\ &\leq \max_{e=(v,w) \in T} \frac{\ell_{\max}}{P(v,w)} \end{aligned}$$

where  $\ell_{\max}$  denotes the maximal length of the canonical paths. If  $\pi$  is not the uniform distribution then a similar estimate holds if the coding is constructed so that

$$(1.8) \quad \pi(v)P(v,w)\pi(\eta_e(x,y)) \approx \pi(x)\pi(y).$$

The *multicommodity flow* technique is a natural extension of canonical paths introduced in Sinclair (1992). Let  $\mathcal{P}_{xy}$  denote the set of all directed paths from  $x$  to  $y$  in the Markov chain on  $\Omega$ . Let  $f_{xy} : \mathcal{P}_{xy} \rightarrow \mathbb{R}_0^+$  for  $x, y \in \Omega$  be functions such that  $\sum_{p \in \mathcal{P}_{xy}} f_{xy}(p) = 1$ . Then each  $f_{xy}$  is called a *1-flow* from  $x$  to  $y$ . In particular any canonical path defines a 1-flow. Then the set  $F = \{f_{xy} \mid x, y \in \Omega\}$  is called a *multicommodity flow*. Sinclair (1992) proved that the estimate of the mixing time in (1.6) extends to multicommodity flows  $F$ ,

$$(1.9) \quad \tau(\varepsilon) \leq \varrho(F)(\log \hat{\pi}^{-1} + \log \varepsilon^{-1})$$

where

$$(1.10) \quad \varrho(F) = \max_{e=(v,w) \in T} \frac{1}{\pi(v)P(v,w)} \sum_{p_{xy} \ni e} \pi(x)\pi(y)f_{xy}(p_{xy})|p_{xy}|$$

is the congestion measure of the flow  $F$ . To bound  $\varrho(F)$  we define, for a transition  $e \in T$  of  $\mathfrak{M}$ , the set  $\mathcal{P}(e) := \bigcup_{x,y \in \Omega} \{p \in \mathcal{P}_{xy} \mid e \in p \text{ and } f_{xy}(p) > 0\}$ . A coding  $(\eta_e)$  for a multicommodity flow  $F$  is a system of (not necessarily injective) mappings  $\eta_e : \mathcal{P}(e) \rightarrow \Omega$  such that for all  $z \in \Omega$ ,

$$(1.11) \quad \sum_{\substack{p_{xy} \in \mathcal{P}(e): \\ \eta_e(p_{xy})=z}} f_{xy}(p_{xy}) \leq 1.$$

Multicommodity flows are in some cases easier to construct than canonical paths. In some recent papers these were instrumental for constructing improved bounds for several basic counting problems (see Sinclair (1992)) and to obtain randomized approximation schemes for long standing open problems like the knapsack problem (Morris and Sinclair (2004)) and for counting contingency tables (Cryan, Dyer, Goldberg, and Jerrum (2002)).

In this paper we first construct and analyse a multicommodity flow for the spanning tree problem and then refine the analysis to canonical paths, which requires some technically involved considerations. In Section 2 we introduce the natural Markov chain for this problem. The mixing time  $\tau_s$  of this chain was bounded from above using coupling arguments by Broder

(1989) and Aldous (1990) who obtained a bound for the mixing time of the order

$$(1.12) \quad \tau_s(\varepsilon) = O(m^2 n^4 (n \log m + \log \varepsilon^{-1}))$$

with  $n = |V|$ ,  $m = |E|$ . Feder and Mihail (1992) improved this bound using the conductance method to the order

$$(1.13) \quad \tau_s(\varepsilon) = O(mn^2 (n \log m + \log \varepsilon^{-1})).$$

In fact they considered an extension of the problem to matroids which satisfy a certain balance condition. Since for a graph  $G = (V, E)$  the pair  $M = (E, \text{ST}(G))$  (where  $\text{ST}(G)$  is the set of spanning trees of  $G$ ) is a graphical matroid satisfying this balance condition, the spanning tree problem is included in their result. In a recent paper Jerrum and Son (2002) found a bound for the log Sobolev constant which leads to an improvement for the Feder and Mihail mixing result for the bases exchange walk for balanced matroids to the order  $O(nm \log n)$ .

Jerrum (1998) suggested that the construction of Cordovil and Moreira (1993) for paths in graphic block matroids was “ideally suited to this purpose” but no further analysis was given in that paper. We prove in detail that based on the paths of Cordovil and Moreira (1993) a multicommodity flow and canonical paths can be defined so that no transition of the Markov chain carries too much weight. This is by no means obvious and needs some careful consideration of the canonical paths (see the proofs of Lemma 3.1 and of Theorem 3.2). As a result we establish that by the method of multicommodity flow one obtains the same bound for the mixing time obtained by Feder and Mihail (1992) by the conductance method but one does not reach the improved bound of Jerrum and Son (2002).

In the final sections we show that the canonical paths for spanning trees are also useful for the analysis of some related Markov chains on the set of forests with roots and on the set of connected spanning subgraphs. In both cases the Markov chains can be shown to be rapidly mixing. Their stationary distribution however is some distribution with weights given by the number of components (for the forest problem) and by the number of spanning trees (for the connected spanning subgraphs problem). For connected subgraphs the interesting case of the uniform distribution remains open.

**2. Markov chain on spanning trees.** In this section we introduce a Markov chain on the set  $\text{ST}(G)$  of spanning trees of an undirected graph  $G = (V, E)$  whose stationary distribution is the uniform distribution on  $\text{ST}(G)$ . We also introduce some notions from graph theory and matroids theory which are used for the construction of multicommodity flows in Section 3.

The Markov chain  $\mathfrak{M}_s(G) = (X_t)_{t \in \mathbb{N}}$  on  $\text{ST}(G)$  is defined by the transition probabilities. If  $X_t = X \in \text{ST}(G)$  is the state of  $\mathfrak{M}_s(G)$  at time  $t \in \mathbb{N}$ , then we draw uniformly and independently  $e \in X$  and  $f \in E$  and set

$$Y = (X \setminus \{e\}) \cup \{f\}.$$

If  $Y \in \text{ST}(G)$  then we set

$$(2.1) \quad X_{t+1} = \begin{cases} Y \\ X \end{cases} \quad \text{each with probability } 1/2.$$

If  $Y \notin \text{ST}(G)$ , then we set  $X_{t+1} = X$ . We denote the transition matrix of  $\mathfrak{M}_s(G)$  by  $P_s$ .

So the transitions of this chain are given by simple random exchanges of two edges as long as they lead again to spanning trees.

**PROPOSITION 2.1.** *The Markov chain  $\mathfrak{M}_s(G)$  is ergodic and the stationary distribution of  $\mathfrak{M}_s(G)$  is the uniform distribution on  $\text{ST}(G)$ .*

*Proof.* Ergodicity is equivalent to irreducibility and aperiodicity. The aperiodicity is obvious by construction, and also  $P_s(X, X) > 1/2$  for any  $X \in \text{ST}(G)$ . For any  $X, Y \in \text{ST}(G)$  we prove by induction on  $k = |X \oplus Y|$ , the cardinality of the symmetric difference of  $X$  and  $Y$ , that  $P_s^t(X, Y) > 0$  for some  $t \in \mathbb{N}$ . Note that  $k = |X \oplus Y| \in 2\mathbb{N}$ . If  $k = 2$ , then  $X \oplus Y = \{a, b\}$  with  $a \in X$  and  $b \in Y$ . Then choosing (with positive probability)  $e = a$  and  $f = b$  in the definition of the chain one gets  $X_{t+1} = Y$  and  $P_s(X, Y) > 0$ . If  $k > 2$  and  $b \in Y \setminus X$  then  $X \cup \{b\}$  contains a circle  $C$  with some edge  $a$  in  $C$  such that  $a \in X \setminus Y$ . Choosing (with positive probability)  $e = a, f = b$ , we obtain  $X' = (X \setminus \{a\}) \cup \{b\} \in \text{ST}(G)$  and  $P_s(X, X') > 0$ . Furthermore,  $|X' \oplus Y| = k - 2$  and so by the induction hypothesis  $P_s^t(X', Y) > 0$  for some  $t \in \mathbb{N}$  and thus  $P_s^{t+1}(X, Y) > 0$ .

For any distinct  $X, Y \in \text{ST}(G)$  with  $P_s(X, Y) > 0$ , we have

$$P_s(X, Y) = \frac{1}{2(n-1)m} = P_s(Y, X)$$

where  $n = |V|, m = |E|$ , i.e. the transition matrix is symmetric and thus reversible with respect to the uniform distribution  $\pi$  on  $\text{ST}(G)$ . This implies that the uniform distribution  $\pi$  is the stationary distribution of  $\mathfrak{M}_s(G)$ . ■

The construction of the multicommodity flow in Section 3 uses for its proof an idea of Cordovil and Moreira (1993) for graphical block-matroids. For ease of reference we recall that a *matroid*  $M = (S, B)$  is a pair of nonempty sets with  $B \subset \mathfrak{P}(S)$  such that for all  $X, Y \in B$  and all  $x \in X \setminus Y$  there exists some  $y \in Y \setminus X$  such that  $(X \setminus \{x\}) \cup \{y\} \in B$ . Elements of  $B$  are called *bases* of  $M$ . A matroid  $M = (S, B)$  is called a *block-matroid* if  $S = X \cup Y$  for some  $X, Y \in B$ .

In our context we consider the graphical matroid  $(E, \text{ST}(G))$  which is a graphical block-matroid if  $G$  can be decomposed into two disjoint spanning trees (see Figure 1).

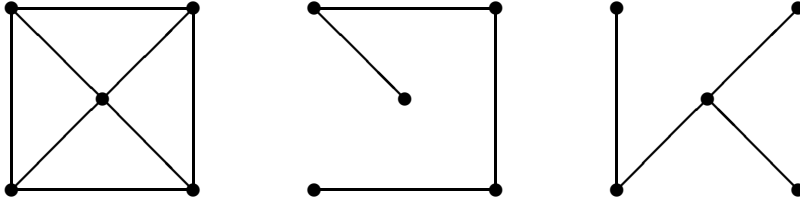


Fig. 1. A graphical block-matroid on the left with a decomposition into two spanning trees

For a block-matroid  $M = (S, B)$  a basis element  $X \in B$  is a *basis-cobasis* of  $M$  if  $S \setminus X \in B$ . The basis-cobasis graph  $H = (V', E')$  is defined by  $V' = \{X \in B \mid X \text{ is a basis-cobasis of } M\}$ ,  $E' = \{\{X_1, X_2\} \subset B \mid |X_1 \oplus X_2| = 2\}$ .

Cordovil and Moreira (1993) proved constructively the following result:

**THEOREM 2.2.** *Let  $M$  be a graphical block-matroid with basis-cobasis graph  $H = (V', E')$ . Then for any  $X, Y \in V'$  there exists a connecting path from  $X$  to  $Y$  in  $H$  of length  $\frac{1}{2}|X \oplus Y|$ .*

We will use the contraction of graphs along edges  $e$  in order to reduce our spanning tree problem to the framework of block-matroids and the basis-cobasis graph. To explain the contraction  $G/e$  on an edge  $e$  see the following example.



Fig. 2. Contraction of a graph on edge  $e$

Here the graph is contracted on edge  $e$  with end nodes  $v, w$ . These contracted nodes give rise to a new node  $v_e$ , while edge  $e$  disappears. For a formal definition see Diestel (1996). For any subset  $S \subset E$  let  $G/S$  denote the contraction on all edges in  $S$ .

**3. Bounding the mixing time of  $\mathfrak{M}_s$ .** We now show how to bound the mixing time of  $\mathfrak{M}_s(G)$  via the multicommodity flow technique described in Section 1.

*Construction of the multicommodity flow  $F_G$ .* For a graph  $G = (V, E)$ , we have to define for each pair  $X, Y \in \text{ST}(G)$  a 1-flow  $f_{XY} : \mathcal{P}_{XY} \rightarrow \mathbb{R}_0^+$  with

$$(3.1) \quad \sum_{p \in \mathcal{P}_{XY}} f_{XY}(p) = 1.$$

To do this, we set  $M := (V, X \cup Y)/(X \cap Y)$  as the graph with the nodes of  $G$  and the edges of  $X \cup Y$  contracted on the edges of  $X \cap Y$ . While the nodes of  $M$  correspond to the connected components of  $(V, X \cap Y)$ , the edges of  $M$  are those of the symmetric difference  $X \oplus Y := (X \cup Y) \setminus (X \cap Y)$ .

Now  $X \setminus Y$  and  $Y \setminus X$  are two disjoint spanning trees of  $M$  and thus  $M$  forms a graphic block-matroid. Each spanning tree of  $M$  can be enhanced to an element of  $\text{ST}(G)$  by adding the edges of  $X \cap Y$ . So the basis-cobasis graph of  $M$  corresponds to a subgraph of the transition graph of  $\mathfrak{M}_s(G)$ . A path  $p' = (B'_i)_{0 \leq i \leq l}$  in this basis-cobasis graph can be transferred to a path  $p = (B_i)_{0 \leq i \leq l}$  in  $\mathfrak{M}_s$  by setting  $B_i := B'_i \cup (X \cap Y)$  for all  $0 \leq i \leq l$ . Exactly to these paths,  $f_{XY}$  will assign a positive weight. The construction due to Cordovil and Moreira (1993) is inductive over the number  $|M|$  of nodes in  $M$ . We next show that this construction can be used to establish inductively a multicommodity flow on  $\text{ST}(G)$ .

If  $|M| = 1$ , then  $X = Y$  and there is nothing to do. In the case  $|M| = 2$ ,  $X$  and  $Y$  are neighbours in  $\mathfrak{M}_s(G)$ . We define  $f_{XY}(p) = 1$  for the path  $p = (B_i)_{0 \leq i \leq 1}$  which consists only of the transition from  $X$  to  $Y$ , i.e.  $B_0 := X$  and  $B_1 := Y$ . For all other  $p' \in \mathcal{P}_{XY} \setminus \{p\}$  we set  $f_{XY}(p') := 0$ , so equation (3.1) holds. Generally we encode a transition  $(B_i, B_{i+1})$  in a path  $(B_i)_{0 \leq i \leq l}$  that carries any weight in  $F_G$  by  $\bar{B}_i := B_0 \oplus B_l \oplus B_i$ . We also have to take care that this encoding is a spanning tree. For the above path  $p$  from  $X$  to  $Y$  we get  $\bar{B}_0 = X \oplus Y \oplus X = Y$ , which is clearly an element of  $\text{ST}(G)$ .

For  $|M| = l + 1$  let  $d_{\min}$  be the minimal degree of a node in  $M$ , and  $D$  the set of nodes of degree  $d_{\min}$ . Then  $d_{\min} = 2$  or  $d_{\min} = 3$  because  $X$  and  $Y$  are spanning trees. Next for each  $v \in D$  we select a pair  $X_v, Y_v \in \text{ST}(G)$  that satisfies the induction hypothesis, and thus  $f_{X_v Y_v}$  is already defined. Further for each path  $p' \in \mathcal{P}_{X_v Y_v}$  with  $f_{X_v Y_v}(p') > 0$  we construct an extension  $p \in \mathcal{P}_{XY}$  and we say  $p$  is *based on*  $p'$ . Let

$$\mathcal{V}(p) := \bigcup_{v \in D} \{p' \in \mathcal{P}_{X_v Y_v} \mid p \text{ is based on } p'\}$$

be the set of all paths  $p$  is based on. Then we define

$$f_{XY}(p) := \sum_{p' \in \mathcal{V}(p)} \frac{1}{|D|} f_{X'Y'}(p'),$$



where  $X'$  and  $Y'$  are start and end nodes of  $p'$  respectively. Since for all  $X, Y \in \text{ST}(G)$ ,

$$\bigcup_{p \in \mathcal{P}_{XY}} \mathcal{V}(p) = \bigcup_{v \in D} \{p' \in \mathcal{P}_{X'Y'} \mid f_{X'Y'}(p') > 0\},$$

the induction hypothesis gives

$$\begin{aligned} \sum_{p \in \mathcal{P}_{XY}} f_{XY}(p) &= \sum_{p \in \mathcal{P}_{XY}} \sum_{p' \in \mathcal{V}(p)} \frac{1}{|D|} f_{X'Y'}(p') = \frac{1}{|D|} \sum_{v \in D} \sum_{p' \in \mathcal{P}_{X_v Y_v}} f_{X_v Y_v}(p') \\ &= \frac{1}{|D|} \sum_{v \in D} 1 = 1. \end{aligned}$$

So  $f_{XY}$  satisfies equation (3.1). We now show how to select for  $v \in D$  the pair  $X_v, Y_v$  and how to derive from the paths in  $\mathcal{P}_{X_v Y_v}$  the paths of  $\mathcal{P}_{XY}$ .

If  $d_{\min} = 2$  let  $a \in X$  and  $b \in Y$  be the two edges in  $M$  at this node  $v$  and set  $X_v := (X \setminus \{a\}) \cup \{b\}$  and  $Y_v := Y$ . For this pair the induction hypothesis holds because  $M_v := (V, X_v \cup Y_v)/(X_v \cap Y_v) = (M/b) \setminus \{a\}$  has exactly  $l$  nodes. A path  $p' = (B'_i)_{0 \leq i < l} \in \mathcal{P}_{X_v Y_v}$  can easily be transformed into a path  $p = (B_i)_{0 \leq i \leq l} \in \mathcal{P}_{XY}$  by adding the transition from  $X$  to  $X_v$  at the first step. Formally  $B_0 := X$  and  $B_{i+1} := B'_i$  for all  $0 \leq i < l$ .

$$(3.2) \quad \begin{array}{ccc} p' : & & X_v = \underbrace{B'_0 \rightarrow \dots \rightarrow B'_{l-1}}_{\downarrow} = Y \\ & & \\ p : & X = B_0 \xrightarrow{-a+b} & \underbrace{B_1 \rightarrow \dots \rightarrow B_l}_{=} = Y. \end{array}$$

The encodings  $(\bar{B}_i)_{0 \leq i < l}$  of  $p$  can also be derived from the encodings  $(\bar{B}'_i)_{0 \leq i < l-1}$  of  $p'$ : We encode the transition  $(B_i, B_{i+1})$  by  $\bar{B}_i := X \oplus Y \oplus B_i = (\bar{B}'_{i+1} \setminus \{b\}) \cup \{a\}$  for  $1 \leq i < l$  and  $\bar{B}_0 := Y$ . These are all spanning trees because in  $\bar{B}'_{i-1}$  the edge  $b$  is the only one at node  $v$ .

Now let  $d_{\min} = 3$  and without loss of generality let  $a, b \in X$  and  $c \in Y$  be the edges at  $v$  in  $M$ . Further let  $a$  be the edge that is included in the only cycle in  $X \cup \{c\}$ . We define  $X_v := X$  and  $Y_v := (Y \setminus \{c\}) \cup \{a\}$  and so the induction hypothesis holds for the pair  $X_v, Y_v$  because the graph  $M_v := (V, X_v \cup Y_v)/(X_v \cap Y_v) = (M/a) \setminus \{c\}$  contains  $l$  nodes.

To derive  $p = (B_i)_{0 \leq i \leq l} \in \mathcal{P}_{XY}$  from a path  $p' = (B'_i)_{0 \leq i < l} \in \mathcal{P}_{X_v Y_v}$  we look at the transition in  $p'$  which exchanges the edge  $b$ , e.g.,  $b \in B'_j \oplus B'_{j+1}$  for some  $j \in \{0, \dots, l-2\}$ . We then define

$$B_i := \begin{cases} B'_i, & i < j, \\ (B'_{i-1} \setminus \{a\}) \cup \{c\}, & i > j, \end{cases}$$

for all  $i \in \{0, \dots, l\} \setminus \{j\}$  and

$$(3.3) \quad B_j := \begin{cases} (B'_{j-1} \setminus \{a\}) \cup \{c\} & \text{if spanning tree,} \\ (B'_{j-1} \setminus \{b\}) \cup \{c\} & \text{otherwise.} \end{cases}$$

$$\begin{array}{ccc} p' : & \underbrace{X \rightarrow \dots \rightarrow B'_{j-1}} & \xrightarrow{-b+d} & \underbrace{B'_j \rightarrow \dots \rightarrow Y_v} \\ & \downarrow & & \downarrow -a+c \\ p : & \underbrace{X \rightarrow \dots \rightarrow B_{j-1}} & \xrightarrow[-b+c]{-a+c} & B_j \xrightarrow[-a+d]{-b+d} & \underbrace{B_{j+1} \rightarrow \dots \rightarrow Y} \end{array}$$

The  $(B_i)_{0 \leq i < j}$  are spanning trees by hypothesis. For  $i > j$  the edge  $a$  is the only edge at  $v$  in  $B'_{i-1}$  and, therefore, replacing  $a$  by  $c$  leads to another spanning tree  $B_i$ . Finally, in (3.3) we remove the edge which is contained in the cycle in  $B'_{i-1} \cup \{c\}$ . This guarantees  $B_j \in \text{ST}(G)$ . Finally, we have to make sure that the encodings  $\bar{B}_i = X \oplus Y \oplus B_i$  are also spanning trees. By definition,

$$\bar{B}_i = \begin{cases} (\bar{B}'_i \setminus \{a\}) \cup \{c\}, & i < j, \\ \bar{B}'_{i-1}, & i > j, \end{cases}$$

for all  $i \in \{0, \dots, l-1\} \setminus \{j\}$ . It follows as before that these are all spanning trees.  $\bar{B}_j = X \oplus Y \oplus B_j$  needs a little more work. In the first case in (3.3) we have

$$\begin{aligned} \bar{B}_j &= X \oplus ((Y_v \setminus \{c\}) \cup \{a\}) \oplus ((B'_{j-1} \setminus \{c\}) \cup \{a\}) \\ &= X \oplus Y_v \oplus B'_{j-1} = \bar{B}'_{j-1}. \end{aligned}$$

Obviously this is a spanning tree by hypothesis. In the second case in (3.3) we have

$$\begin{aligned} \bar{B}_j &= X \oplus ((Y_v \setminus \{c\}) \cup \{a\}) \oplus ((B'_{j-1} \setminus \{c\}) \cup \{b\}) \\ &= X \oplus Y_v \oplus B'_{j-1} \oplus \{a, b\} = (\bar{B}'_{j-1} \setminus \{a\}) \cup \{b\}. \end{aligned}$$

The cycle in  $\bar{B}'_{j-1} \cup \{b\}$  must include  $a$  and therefore also in this case  $\bar{B}_j$  is a spanning tree.

The multicommodity flow  $F_G := \{f_{XY} \mid X, Y \in \text{ST}(G)\}$  on  $\mathfrak{M}_s(G)$  and its codings are now defined. Because the functions  $f_{XY} : \mathcal{P}_{XY} \rightarrow \mathbb{R}_0^+$  are defined on disjoint sets we can look at  $F_G$  as a function from  $\mathcal{P} := \bigcup_{X, Y \in \text{ST}(G)} \mathcal{P}_{XY}$  to  $\mathbb{R}_0^+$  and write

$$F_G(p) = f_{XY}(p) \quad \text{for all } p \in \mathcal{P}_{XY} \text{ and all } X, Y \in \text{ST}(G).$$

The inductive construction of  $F_G$  has the positive effect that a 1-flow  $f_{XY} \in F_G$  does not differ much from another flow  $f_{X'Y'} \in F_G$  if the spanning trees  $X, X'$  and  $Y, Y'$  are very similar. The following lemma makes this clear.

LEMMA 3.1. Given  $X, Y \in \text{ST}(G)$  and a node  $w$  in the graph  $M := (V, X \cup Y)/(X \cap Y)$  of degree 3. The three edges in  $M$  at  $w$  are  $a, b \in Y$  and  $c \in X$ . Then  $X^a := (X \setminus \{c\}) \cup \{a\}$  and  $X^b := (X \setminus \{c\}) \cup \{b\}$  are also spanning trees. Furthermore, let  $p_a = (A_i)_{0 \leq i \leq l} \in \mathcal{P}_{X^a Y}$  and  $p_b = (B_i)_{0 \leq i \leq l} \in \mathcal{P}_{X^b Y}$  be the paths with

$$(3.4) \quad A_i = \begin{cases} B_i \oplus \{a, b\}, & 0 \leq i \leq j, \\ B_i, & j < i \leq l, \end{cases}$$

where  $j$  is the step in  $p_a$  which exchanges  $b$ , i.e.  $b \in A_j \oplus A_{j+1}$ . Then for these paths  $p_a, p_b$  and the multicommodity flow  $F_G$  we have

$$F_G(p_a) = F_G(p_b).$$

Similarly, for  $p'_a \in (A'_i)_{0 \leq i \leq l} \in \mathcal{P}_{Y X^a}$  and  $p'_b = (B'_i)_{0 \leq i \leq l} \in \mathcal{P}_{Y X^b}$  with

$$A'_i = \begin{cases} B'_i, & 0 \leq i \leq j, \\ B'_i \oplus \{a, b\}, & j < i \leq l, \end{cases}$$

where again  $j \in \{0, \dots, l-1\}$  with  $b \in A'_j \oplus A'_{j+1}$ , we have  $F_G(p'_a) = F_G(p'_b)$ .

*Proof.* By definition of  $X^a$  and  $X^b$ , one can easily see that the graphs  $M^a := (V, X^a \cup Y)/(X^a \cap Y)$  and  $M^b := (V, X^b \cup Y)/(X^b \cap Y)$  are isomorphic. If  $a$  connects the nodes  $w$  and  $t$ , and  $b$  the nodes  $w$  and  $s$  in  $M$ , then  $I : M^a \rightarrow M^b$  given by

$$I(v) := \begin{cases} v, & v \notin \{v_a, s\}, \\ t, & v = v_a, \\ v_b, & v = s, \end{cases} \quad I(e) := \begin{cases} e, & e \neq b, \\ a, & e = b, \end{cases}$$

for all nodes  $v$  and edges  $e$  in  $M^a$  defines an isomorphism. Figure 3 shows an example.

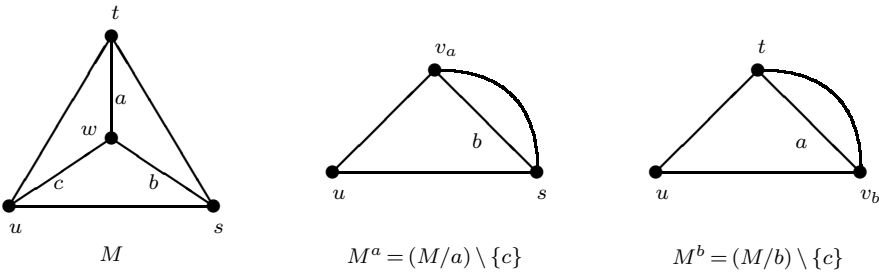


Fig. 3. An example for a graph  $M$  from Lemma 3.1. The resulting graphs  $M^a$  and  $M^b$  are isomorphic.

The numbers of nodes in  $M^a$  and in  $M^b$  are the same. The proof of Lemma 3.1 is by induction on  $|M^a|$ :

For  $|M^a| = 2$  the path  $p_a = (A_i)_{0 \leq i \leq 1}$  given by  $A_0 = X^a$  and  $A_1 = Y$  is the only path in  $\mathcal{P}_{X^a Y}$  with a positive weight in  $F_G$ , i.e.  $F_G(p_a) = 1$ . The

same holds for  $p_b = (B_i)_{0 \leq i \leq l}$  given by  $B_0 = X^b = (A_0 \setminus \{a\}) \cup \{b\}$  and  $B_l = Y = A_l$ . So in this case the lemma is proved.

If  $|M^a| = l + 1$  let  $D_a$  be the set of nodes in  $M^a$  of minimal degree  $d_{\min}$  and  $D_b$  the analogous set of nodes in  $M^b$ . Each node in  $D_a$  corresponds to a node in  $D_b$  via the isomorphism  $I$ .

The value of  $F_G(p_a)$  is based by construction on the paths in  $\mathcal{V}(p_a)$ . For  $v \in D_a$  each path  $p'_a \in \mathcal{P}_{X_v^a Y_v} \cap \mathcal{V}(p_a)$  corresponds to a path  $p'_b \in \mathcal{P}_{X_u^b Y_u} \cap \mathcal{V}(p_b)$ ,  $u := I(v) \in D_b$ . If  $a$  is an edge at  $v$  in  $M^a$ , then  $b$  is an edge at  $u$  in  $M^b$ . This leads to  $X_v^a = X_u^b$  and  $Y_v = Y_u$  and, therefore,  $p'_a = p'_b$ . Otherwise, if  $a$  is not an edge at  $v$  we have  $v = u$ . In this case either  $X^a$  arises from  $X_v^a$  and  $Y_v = Y$ , or  $X_v^a = X^a$  and  $Y$  arises from  $Y_v$  by exchanging two edges. The same holds for  $X_v^b$  and  $Y_v$ . This modification can also be done at  $X$  and  $Y$ , and for the resulting spanning trees  $X_v, Y_v$  and the node  $w$  the induction hypothesis holds for all paths in  $\mathcal{P}_{X_v^a Y_v}$  and  $\mathcal{P}_{X_v^b Y_v}$  that satisfy (3.4). This condition is in particular satisfied by  $p'_a$  and  $p'_b$  and, therefore,  $F_G(p'_a) = F_G(p'_b)$ . It follows that  $F_G(p'_a) = F_G(p'_b)$  for all  $p'_a \in \mathcal{V}(p_a)$  and for the corresponding  $p'_b \in \mathcal{V}(p_b)$ . Finally,

$$F_G(p_a) = \frac{1}{|D_a|} \sum_{p'_a \in \mathcal{V}(p_a)} F_G(p'_a) = \frac{1}{|D_b|} \sum_{p'_b \in \mathcal{V}(p_b)} F_G(p'_b) = F_G(p_b). \blacksquare$$

The multicommodity flow  $F_G$  yields good upper bounds for the mixing time of  $\mathfrak{M}_s(G)$  if no transition of the Markov chain carries too much weight. The following theorem shows that this holds for  $F_G$ .

**THEOREM 3.2.** *Let  $B$  and  $\bar{B}$  be two spanning trees of a graph  $G = (V, E)$  and let  $e \in B \oplus \bar{B}$  be an edge. Define  $\mathcal{B} := \mathcal{B}(B, \bar{B}, e)$  as the set of paths  $p = (B_i)_{0 \leq i \leq l} \in \mathcal{P}$  such that there exists  $j \in \{0, \dots, l - 1\}$  with*

- (a)  $B = B_j$ , i.e.,  $p$  contains  $B$ ,
- (b)  $\bar{B} = B_0 \oplus B_l \oplus B$ , i.e., the coding of  $B$  in  $p$  is  $\bar{B}$ ,
- (c)  $e \in B_j \oplus B_{j+1}$ , i.e.,  $p$  leaves  $B$  by exchanging  $e$ ,
- (d)  $F_G(p) > 0$ .

Then

$$\sum_{p \in \mathcal{B}} F_G(p) = 1.$$

*Proof.* We set  $M := (V, B \cup \bar{B}) / (B \cap \bar{B})$  and again we proceed inductively on  $|M|$ . For  $|M| = 2$  the set  $B \oplus \bar{B}$  contains exactly two edges. Thus  $\mathcal{B}$  contains only the path  $p = (B_i)_{0 \leq i \leq l}$  with  $B_0 = B$  and  $B_l = \bar{B}$ . By definition  $F_G(p) = 1$ .

Now let  $|M| = l + 1$  and  $D$  be the set of nodes in  $M$  of minimal degree  $d_{\min}$ . For each  $v \in D$  we construct a set  $\mathcal{B}_v \subset \mathcal{P}$  with

- (i)  $\sum_{p' \in \mathcal{B}_v} F_G(p') = 1,$
- (ii)  $\sum_{p \in \mathcal{B}} F_G(p) = \sum_{v \in D} \sum_{p' \in \mathcal{B}_v} \frac{1}{|D|} F_G(p').$

The theorem then follows easily, because

$$\sum_{p \in \mathcal{B}} F_G(p) = \sum_{v \in D} \sum_{p' \in \mathcal{B}_v} \frac{1}{|D|} F_G(p') = \frac{1}{|D|} \sum_{v \in D} 1 = 1.$$

If  $d_{\min} = 2$ , let  $b \in B$  and  $a \in \bar{B}$  be the two edges at a fixed  $v \in D$ . If  $e \notin \{a, b\}$ , we define  $B_v := B$  and  $\bar{B}_v := (\bar{B} \setminus \{a\}) \cup \{b\}$ . The graph  $M_v := (V, B_v \cup \bar{B}_v) / (B_v \cap \bar{B}_v)$  has  $l$  nodes. Therefore, properties (a)–(d) and also (i) are satisfied by the set  $\mathcal{B}_v := \mathcal{B}(B_v, \bar{B}_v, e)$ . If  $e \in \{a, b\}$ , we set  $B_v := (B \setminus \{b\}) \cup \{a\}$  and  $\bar{B}_v := B_v$ . Then  $\mathcal{B}_v := \mathcal{P}_{B_v \bar{B}_v}$  satisfies (i) by definition of  $F_G$ . To ensure (ii), we show

$$(3.5) \quad \bigcup_{p \in \mathcal{B}} \mathcal{V}(p) = \bigcup_{w \in D} \mathcal{B}_w.$$

This gives (ii) immediately:

$$\sum_{p \in \mathcal{B}} F_G(p) = \sum_{p \in \mathcal{B}} \sum_{p' \in \mathcal{V}(p)} \frac{1}{|D|} F_G(p') = \sum_{v \in D} \sum_{p' \in \mathcal{B}_v} \frac{1}{|D|} F_G(p').$$

While in (3.5),  $\bigcup \mathcal{V}(p) \subseteq \bigcup \mathcal{B}_w$  is a consequence of the construction of  $F_G$ , the inclusion  $\bigcup \mathcal{V}(p) \supseteq \bigcup \mathcal{B}_w$  is obtained as follows: Take  $p \in \mathcal{B}$  and  $p' \in \mathcal{V}(p)$ , which is based on  $p$  via a node  $w \in D$ . If  $e \notin \{a, b\}$ , then  $p'$  is extended to  $p$  by adding the exchange of  $a$  and  $b$  at the first step. By definition of  $\mathcal{B}$ ,  $p$  leaves the spanning tree  $B = B_v$  by exchanging  $e$ . The same holds, therefore, for  $p'$ . But in  $p'$  this transition is coded by  $\bar{B}' := (\bar{B} \setminus \{a\}) \cup \{b\}$  and so  $p' \in \mathcal{B}(B_v, \bar{B}_v, e) = \mathcal{B}_v$ . If  $e \in \{a, b\}$  we get  $p \in \mathcal{P}_{B \bar{B}}$ , because  $a, b$  are exchanged first in  $p$ . Thus the path  $p'$  is a path from  $B' := (B \setminus \{b\}) \cup \{a\} = B_v$  to  $\bar{B} = B_v$  and hence  $p' \in \mathcal{B}_v = \mathcal{P}_{B_v \bar{B}_v}$ .

The case  $d_{\min} = 3$  is treated analogously. For  $v \in D$  let (without loss of generality)  $a, b \in B$  and  $c \in \bar{B}$  be the tree edges at  $v$  in  $M$ . We further define  $B_v := B$ ,  $\bar{B}_v^a := (\bar{B} \setminus \{c\}) \cup \{a\}$  and  $\bar{B}_v^b := (\bar{B} \setminus \{c\}) \cup \{b\}$ . Let  $X, Y \in \text{ST}(G)$  be the start and end nodes of a path  $p \in \mathcal{B}$  that is based on  $p' \in \mathcal{V}(p)$  via a node  $w \in D$ . We first consider  $e \notin \{a, b, c\}$  and again  $a, b \in X$ ,  $c \in Y$ . According to the construction of  $F_G$ , the edges  $a, b$ , and  $c$  are exchanged in  $p$  in two consecutive transitions. Because  $a, b \in B$ , both edges are either in  $X$  or in  $Y$ . The only cycle in  $X \cup \{c\}$  contains either  $a$  or  $b$ . If it contains  $a$  then  $p' \in \mathcal{B}^a := (B_v, \bar{B}_v^a, e)$ , else  $p' \in \mathcal{B}^b := (B_v, \bar{B}_v^b, e)$ . For  $p' = (B_i)_{0 \leq i \leq l} \in \mathcal{B}^a$ ,

we define  $p'' = (A_i)_{0 \leq i \leq l} \in \mathcal{B}^b$  by

$$A_i := \begin{cases} B_i, & 0 \leq i \leq j, \\ B_i \oplus \{a, b\}, & j < i \leq l, \end{cases}$$

with  $j \in \{0, \dots, l - 1\}$  such that  $b \in B_j \oplus B_{j+1}$ . These paths  $p', p''$ , the spanning trees  $X, Y$  and the node  $w$  of degree 3 satisfy the assumptions of Lemma 3.1 and hence  $F_G(p') = F_G(p'')$ . Furthermore,  $p''$  cannot be in any other  $\mathcal{V}(\tilde{p})$ . If this were the case, then  $p \in \mathcal{P}_{XY}$  and so  $\tilde{p} = p$ . For any  $p \in \mathcal{B}$ , the path  $p' \in \mathcal{V}(p)$  based on  $w \in D$  is contained either in  $\mathcal{B}^a$  or in  $\mathcal{B}^b$ , while the corresponding  $p''$  cannot be in another set  $\mathcal{V}(\tilde{p})$ . Thus  $F_G(p') = F_G(p'')$  and by the induction hypothesis

$$\sum_{p' \in \mathcal{B}^a} F_G(p') = \sum_{p' \in \mathcal{B}^b} F_G(p') = 1.$$

The set

$$(3.6) \quad \mathcal{B}_v := \{p' \in \mathcal{B}^a \cup \mathcal{B}^b \mid \exists \tilde{p} \in \mathcal{B} : p' \in \mathcal{V}(\tilde{p})\}$$

satisfies (i):

$$\sum_{p' \in \mathcal{B}_v} F_G(p') = \sum_{p' \in \mathcal{B}^a} F_G(p') = \sum_{p' \in \mathcal{B}^b} F_G(p') = 1.$$

If  $e \in \{a, b, c\}$ , then we set  $\mathcal{B}^a := \mathcal{B}(B_v, \bar{B}_v^a, b)$  and  $\mathcal{B}^b := \mathcal{B}(B_v, \bar{B}_v^b, a)$  and proceed as above. In this case we also get a set  $\mathcal{B}_v$  with property (i). The definition (3.6) of  $\mathcal{B}_v$  ensures that

$$\bigcup_{p \in \mathcal{B}} \mathcal{V}(p) = \bigcup_{v \in D} \mathcal{B}_v$$

and, therefore,  $\mathcal{B}_v$  satisfies (ii). Now the theorem follows as in the case  $d_{\min} = 2$ . ■

It is not difficult to see that Theorem 3.2 guarantees (1.11). All preparations are now made to obtain an efficient upper bound on the mixing time of  $\mathfrak{M}_s(G)$ .

**THEOREM 3.3.** *For a graph  $G = (V, E)$  the mixing time  $\tau_s$  of the Markov chain  $\mathfrak{M}_s(G)$  is bounded by*

$$\tau_s(\varepsilon) \leq 2n^2m(n \log m + \log \varepsilon^{-1})$$

for all  $\varepsilon \in (0, 1)$  with  $n := |V|$  and  $m := |E|$ .

*Proof.* In Proposition 2.1 we have seen that the Markov chain  $\mathfrak{M}_s(G)$  meets the prerequisites of the result of Diaconis and Stroock (1991) mentioned in (1.6) and its extension by Sinclair (1992) in (1.12). Hence we already have

$$(3.7) \quad \tau_s(\varepsilon) \leq \varrho(F_G)(\log \hat{\pi}^{-1} + \log \varepsilon^{-1})$$

with  $\hat{\pi} := \min_{x \in \text{ST}(G)} \pi(x)$  and

$$\varrho(F_G) := \max_{\substack{v,w \in \Omega \\ P_s(v,w) > 0}} \frac{1}{\pi(v)P_s(v,w)} \sum_{p_{xy} \in \mathcal{P}(v,w)} \pi(x)\pi(y)f_{xy}(p_{xy})|p_{xy}|,$$

where  $\mathcal{P}(v, w)$  is the set of paths  $p$  that contain the transition  $(v, w)$  and have  $F_G(p) > 0$ . By construction the length of path  $p$  with positive weight in  $F_G$  is at most  $n-1$ , because two spanning trees differ in at most  $2(n-1)$  edges. Furthermore the stationary distribution of  $\mathfrak{M}_s(G)$  is the uniform distribution on  $\text{ST}(G)$  and the transition probabilities are either  $1/2m(n-1)$  or 0.

This gives

$$(3.8) \quad \varrho(F_G) \leq \frac{2n^2m}{|\text{ST}(G)|} \max_{\substack{v,w \in \Omega \\ P_s(v,w) > 0}} \sum_{p \in \mathcal{P}(v,w)} F_G(p).$$

We now use Theorem 2.2 to bound the second factor of this estimate: For an arbitrary transition  $(v, w)$  of  $\mathfrak{M}_s(G)$  and  $p = (B_i)_{0 \leq i \leq l} \in \mathcal{P}(v, w)$  there is some  $j \in \{0, \dots, l-1\}$  with  $v = B_j$ . Let this  $v$  be encoded by  $\bar{v}$ . As  $v \setminus w$  contains exactly one edge  $e$  we obtain  $p \in \mathcal{B}(v, \bar{v}, e)$  and hence

$$\mathcal{P}(v, w) \subset \bigcup_{\bar{v} \in \text{ST}(G)} \mathcal{B}(v, \bar{v}, e).$$

From Theorem 2.2 we deduce

$$\sum_{p \in \mathcal{P}(v,w)} F_G(p) \leq \sum_{\bar{v} \in \text{ST}(G)} \sum_{p \in \mathcal{B}(v,\bar{v},e)} F_G(p) = |\Omega|,$$

and by (3.8),

$$\varrho(F_G) \leq \frac{2n^2m}{|\text{ST}(G)|} |\text{ST}(G)| = 2n^2m.$$

Together with the rough bound  $|\text{ST}(G)| \leq m^n$ , from (3.7) we finally get for all  $\varepsilon \in (0, 1)$ ,

$$\tau_s(\varepsilon) \leq 2n^2m(n \log m + \log \varepsilon^{-1}). \blacksquare$$

**4. Canonical paths for  $\mathfrak{M}_s$ .** For the construction of canonical paths for  $\mathfrak{M}_s(G)$  we shall make use of the multicommodity flow  $F_G$  of Section 3. For  $X, Y \in \text{ST}(G)$  for the construction of a 1-flow  $f_{XY}$  for any node  $v$  of minimal degree in the contracted graph  $M = (V, X \cup Y)/(X \cap Y)$  we used by induction a 1-flow  $f_{X_v Y_v}$  already constructed. If in this recursion this node  $v$  is always uniquely determined, then  $f_{XY}$  is in fact a 1-flow along some path in  $\mathfrak{M}_s(G)$  since the construction begins with a simple transition between neighbours. To obtain canonical paths for  $\mathfrak{M}_s(G)$  we have to determine which of the nodes of minimal degree has to be chosen in the recursion step. In the following, we call this node the *starting node* in  $M$ .

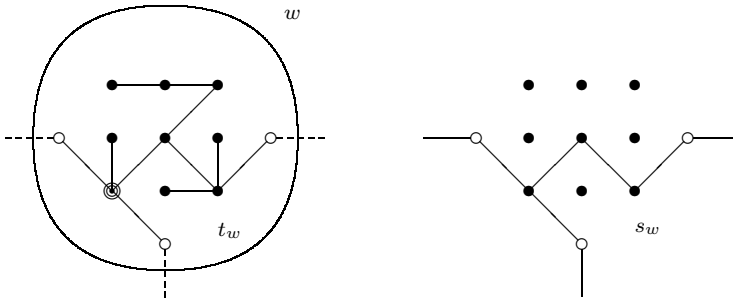


Fig. 4. Boundary nodes in  $t_w = 0$ , boundary edges = - - -, internal node  $\odot$  is of degree 3 in  $s_w$

To construct the starting node in  $M$  we assume without loss of generality that  $M$  has at least three nodes. If there is exactly one node of minimal degree we call it the starting node. In the other case we number the nodes in  $V$  by  $1, \dots, n$  and consider the subgraph  $M' := (V, X \cup Y)$ . We map each node with minimal degree in  $M$  injectively to the index of a node in  $V$  and choose as starting node of  $M$  the one of minimal index. A node  $w$  in  $M$  corresponds to a connected component of  $X \cap Y$  and thus to a subtree  $t_w$  of  $M'$ . A node of  $t_w$  is called a *boundary node* if it is an end node of edges in  $M'$  which are not in  $X \cap Y$  but in  $X \oplus Y$ . These edges are called *boundary edges*. Any pair of nodes in the tree  $t_w$  is connected by exactly one path in  $t_w$ . We denote by  $s_w$  the subgraph of  $M'$  consisting of the paths which connect boundary node pairs, supplemented by boundary edges. A node in  $t_w$  is called an *internal node* if its degree in  $s_w$  is  $\geq 3$ .

If  $w$  in  $M$  has minimal degree  $d_{\min} = 3$ , then  $t_w$  has exactly one internal node whose index we attach to  $w$ . This can be seen as follows. If  $t_w$  has only one boundary node  $p$ , then the boundary edges are the only edges in  $s_w$  and so  $p$  is the only node from  $t_w$  in  $s_w$  of degree 3. If there are two boundary nodes  $p$  and  $q$ , then  $s_w$  consists of the path between  $p$  and  $q$  and the three boundary edges. In  $s_w$  there is only one boundary node of degree 3 with two boundary edges. In case  $t_w$  has three different boundary nodes  $p, q$  and  $r$  in the first case the path in  $t_w$  between two of these nodes might contain the third one. If e.g.  $r$  is on the path from  $p$  to  $q$ , then only  $r$  can have degree 3 in  $s_w$  induced by the boundary edge at this node. If no boundary node is on the path between two others, then the final segments of the path from  $q$  to  $p$  and from  $r$  to  $p$  coincide. Let  $u$  denote the end node of this segment other than  $p$ ; then  $s_w$  consists of three disjoint paths to the boundary nodes starting at  $u$ , and  $u$  is the only node from  $t_w$  in  $s_w$  of degree 3 (see Figure 4). Since  $w$  is thus attached the index of a node in  $t_w$ , this mapping to the index is injective.



If the minimal degree  $d_{\min}$  in  $M$  is 2, then for no node  $w$  of degree 2 does the corresponding subtree  $t_w$  in  $M'$  have an inner node, since it has at most two boundary nodes. Further, those  $w$  are not neighbours to further nodes of degree 2 as  $M$  is the union of two disjoint spanning trees. We now attach to both edges at  $w$  in  $M$  a node in  $V$  and then choose as partner of  $w$  the smaller of these two indices. The following procedure is demonstrated in Figures 5 and 6: Let  $e$  be an edge in  $M$  connecting  $w$  with a node  $v$ . Then  $e$  is also in  $M'$  and connects there a boundary node  $w'$  of  $t_w$  with a boundary node  $v'$  in  $t_v$ . If  $v'$  is an inner node of  $t_v$  then we attach to  $e$  the node  $w$ . In the other case we consider the subgraph  $s_v$  where  $v'$  has degree 2 since it is a boundary node of  $t_v$  but it is not an inner node. Therefore, by construction all other boundary nodes are connected with  $v'$  by a path in  $s_v$  and coincide on the initial segment from  $v'$  to some inner point  $v''$ . We attach to  $e$  not the inner node  $v''$  but its neighbour in this segment. As a result our mapping is injective.

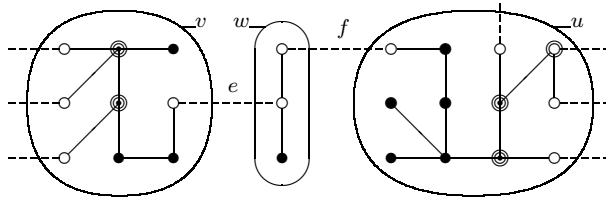


Fig. 5. In the middle part we have node  $w$  connected via  $e, f$  with  $v$  and  $u$  in  $M$ . Boundary nodes =  $\circ$ , inner nodes =  $\odot$ .

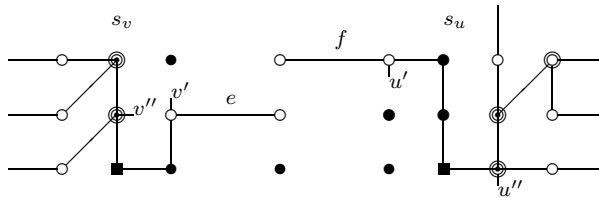


Fig. 6. Corresponding subgraphs  $s_v$  and  $s_u$ . Determining the index of node  $w$  of degree 2. In  $s_v$  all paths from  $v'$  to boundary nodes coincide on the initial edge part up to  $v''$ . Neighbour of  $v'' = \blacksquare$ . Analogously in  $s_u$ . The smaller index of the  $\blacksquare$ -node is attached to  $w$ .

After this involved determination of the starting point we can follow the construction of multicommodity flows of Section 3 and construct canonical paths for the Markov chain  $\mathfrak{M}_s$ .

*Construction of canonical paths in  $\mathfrak{M}_s(G)$ .* For  $X, Y \in \text{ST}(G)$  and  $M := (V, X \cup Y)/(X \cap Y)$  let  $F_G$  be the multicommodity flow of  $\mathfrak{M}_s(G)$  as in Section 3. The canonical path  $\gamma_{XY}$  from  $X$  to  $Y$  in  $\mathfrak{M}_s(G)$  is defined by

induction on  $|M|$ : If  $|M| = 2$  then  $X$  and  $Y$  are neighbours and the canonical path is  $\gamma_{XY} = (B_i)_{0 \leq i \leq 1}$ ,  $B_0 := X$ ,  $B_1 = Y$ . Since  $F_G(\gamma_{XY}) = 1$  we choose the corresponding coding  $\bar{B}_0 := Y$ ,  $\bar{B}_1 := X$ . For the induction step  $|M| = l + 1$  we proceed as in the construction of  $F_G$ . There, however, we determined for any node  $v$  the set  $D$  of all nodes of minimal degree in  $M$ ,  $X_v, Y_v \in \text{ST}(G)$ , we constructed for any path  $p' \in \mathcal{P}_{X_v Y_v}$  with  $F_G(p') > 0$  a path  $p \in \mathcal{P}_{XY}$ , and we determined  $f_{XY}(p)$  as the sum of all  $f_{X_v Y_v}(p')$  over all  $v \in D$  and  $p' \in \mathcal{P}_{X_v Y_v}$  normalized by  $1/|D|$ . Now by induction hypothesis, for the starting node  $v_0$  of  $M$  and the corresponding  $X' := X_{v_0}$ ,  $Y' := Y_{v_0}$ , there already exists exactly one canonical path  $\gamma_{X'Y'}$ . As in the construction of  $p \in \mathcal{P}_{XY}$  from  $p' \in \mathcal{P}_{X_v Y_v}$  we obtain a canonical path  $\gamma_{XY}$  from  $\gamma_{X'Y'}$  and the coding of a transition  $(B_i, B_{i+1})$  in  $\gamma_{XY}$  is given in the same way by  $\bar{B}_i := X \oplus Y \oplus B_i$ . We denote by  $\Gamma_G$  the set of all canonical paths in  $\mathfrak{M}_s(G)$ .

We next obtain results analogous to Lemma 3.1 and Theorem 3.2. The canonical paths  $\gamma_{XY}$  and  $\gamma_{X'Y'}$  are similar if the pairs  $X, Y$  and  $X', Y'$  are close.

LEMMA 4.1. *Let  $X, Y \in \text{ST}(G)$  and  $w$  a node of degree 3 in  $M$ . If  $a, b \in Y$  and  $c \in X$  are edges to  $w$  in  $M$ , then  $X^a := (X - \{c\}) \cup \{a\}$  and  $X^b := (X - \{c\}) \cup \{b\}$  are in  $\text{ST}(G)$  and for the canonical paths  $\gamma_{X^a Y} := (A_i)_{0 \leq i \leq l}$  and  $\gamma_{X^b Y} := (B_i)_{0 \leq i \leq l}$  we have*

$$A_i = \begin{cases} B_i \oplus \{a, b\}, & 0 \leq i \leq k, \\ B_i, & k < i \leq l, \end{cases}$$

where  $k$  is the place in  $\gamma_{X^a Y}$  where the edge  $b$  is added, i.e.  $b \in A_k \oplus A_{k+1}$ . Similarly for  $\gamma_{Y X^a} := (A'_i)_{0 \leq i \leq l}$ ,  $\gamma_{Y X^b} := (B'_i)_{0 \leq i \leq l}$  we have

$$A'_i = \begin{cases} B'_i, & 0 \leq i \leq k \text{ with } k \in \{0, \dots, l-1\}, \\ B_i \oplus \{a, b\}, & k < i \leq l, \end{cases}$$

so that  $b \in A'_k \oplus A'_{k+1}$ .

The injectivity of the coding is needed to prove that no transitions are used in too many canonical paths. This is a consequence of the following theorem which is parallel to Theorem 3.2.

THEOREM 4.2. *For  $B, \bar{B} \in \text{ST}(G)$  and edge  $e \in B \oplus \bar{B}$  there exists exactly one canonical path  $\gamma_{X,Y} := (B_i)_{1 \leq i \leq l}$  and some  $j \in \{0, \dots, l-1\}$  such that*

- (a)  $B = B_j$ , i.e.  $\gamma_{XY}$  contains  $B$ ,
- (b)  $\bar{B} = X \oplus Y \oplus B$ , i.e.  $B$  is coded in  $\gamma_{XY}$  by  $\bar{B}$ ,
- (c)  $e \in B_j \oplus B_{j+1}$ , i.e. in  $\gamma_{XY}$  directly after  $B$  the edge  $e$  is replaced.

Detailed proofs of Lemma 4.1 and Theorem 4.2 are given in Fehrenbach (2003). As a consequence we obtain a proof of the mixing time bound in Theorem 3.3 by using canonical paths:

*Proof of Theorem 3.3 by using canonical paths.* For the mixing time  $\tau = \tau(\varepsilon)$ , by (1.5) we have

$$(4.1) \quad \tau(\varepsilon) \leq \varrho(\Gamma_G)(\log \hat{\pi} + \log \varepsilon^{-1})$$

where  $\pi$  is the stationary distribution,  $\hat{\pi} := \min_{x \in \Omega} \pi(x)$  and

$$(4.2) \quad \varrho(\Gamma_G) := \max_{\substack{(B,C) \in \Omega^2 \\ P_s(B,C) > 0}} \frac{1}{\pi(B) P_s(B,C)} \sum_{\gamma_{XY} \in \mathcal{P}(B,C)} \pi(X)\pi(Y)|\gamma_{XY}|,$$

$\mathcal{P}(B, C)$  being the set of canonical paths which contain the transition  $(B, C)$ . The maximal length of a canonical path in  $\Gamma_G$  is  $n - 1$  since at most  $2(n - 1)$  edges are exchanged. Further  $P_s(B, C) = 1/2(n - 1)m$ ,  $m = |E|$ . Thus for the congestion measure we get

$$(4.3) \quad \varrho(\Gamma_G) \leq \frac{2n^2m}{|\Omega|} \max_{\substack{(B,C) \in \Omega^2 \\ P_s(B,C) > 0}} |\mathcal{P}(B, C)|.$$

The max expression can be estimated using Theorem 4.2. For a transition  $(B, C)$  of  $\mathfrak{M}_s(G)$  let  $e$  be the unique edge in  $B \setminus C$ . In any canonical path  $\gamma_{XY} \in \mathcal{P}(B, C)$  the transition  $(B, C)$  is coded by some  $\bar{B} \in \Omega$ . By Theorem 4.2 there exists exactly one canonical path in  $\Gamma_G$  with these properties. Since this path not necessarily contains  $(B, C)$  we conclude that  $|\mathcal{P}(B, C)| \leq |\Omega|$ . This implies

$$(4.4) \quad \varrho(\Gamma_G) \leq \frac{2n^2m}{|\Omega|} |\Omega| = 2n^2m.$$

Since  $|\Omega| \leq m^n$  we thus obtain

$$(4.5) \quad \tau_s(\varepsilon) \leq 2n^2m(n \log m + \log \varepsilon^{-1}).$$

**5. Forests with roots.** In this section we apply the multicommodity flows and canonical paths to the analysis of the mixing time of some Markov chains on forests. The Markov chain  $\mathfrak{M}_s$  introduced in Sections 2 and 3 on the set of spanning trees only uses exchanges of two edges. These transitions can also be used on the class of forests, i.e. cycle-free subgraphs of  $G$ , and the corresponding Markov chain has as stationary distribution the uniform distribution. But so far no efficient bounds for the mixing time of this or related Markov chains are known and also no randomized approximation schemes for the number of forests are known (see Welsh and Merino (2000)). It seems that also the canonical paths of Section 4 transferred to this problem do not lead to a polynomial bound for the mixing time.

In the following we consider a modified class  $F_r(G)$  of forests with roots and show that for this modified space  $\Omega = F_r(G)$  we obtain rapid mixing

results for various Markov chains by means of the corresponding canonical paths constructed for the class of spanning trees.

DEFINITION 5.1. Let  $G = (V, E)$  be an undirected graph. A pair  $X := (R_X, E_X)$  with  $R_X \subset V$  and  $E_X \subset E$  is called a *forest with roots* if

- the subgraph  $(V, E_X)$  of  $G$  contains no cycle,
- any connected component  $Z$  of  $(V, E_X)$  has exactly one node in  $R_X$ , which we call the root of  $Z$ .

$F_r(G)$  denotes the set of all forests with roots.

Counting forests with roots corresponds to counting forests  $X$  with connected components  $Z_1, \dots, Z_d$  which are weighted by the number of possibilities to choose a root system, i.e. by  $\prod_{i=1}^d |Z_i|(n - |X|)$ . The class  $F_r(G)$  of forests with roots can be identified with the class of spanning trees of an extended graph  $G'$ .

LEMMA 5.2. *For any undirected graph  $G = (V, E)$  there exists a graph  $G' = (V', E')$  such that there exists a bijection  $\text{Sp} : F_r(G) \rightarrow \text{ST}(G')$ .*

*Proof.* Let  $V' := V \cup \{r\}$  and  $E' := E \cup \{\{r, v\} \mid v \in V\}$ , i.e. we add a node  $r$  and all edges  $\{r, v\}$  to the new node to obtain  $G' = (V', E')$ . For  $X \in F_r(G)$  we define  $\text{Sp}(X) := E_X \cup \{\{r, v\} \mid v \in R_X\}$ . Thus  $\text{Sp}(X)$  is a spanning tree of  $G'$  since  $E_X$  is cycle-free and also the addition of the edges  $\{r, v\}$ ,  $v \in R_X$ , does not produce cycles.  $\text{Sp}$  is a bijection since any  $X' \in \text{ST}(G')$  has a unique origin  $\text{Sp}^{-1}(X') = (R_X, E_X)$  where  $R_X := \{v \in V \mid \{r, v\} \in X'\}$  and  $E_X := X' \cap E$ . ■

The bijection of Lemma 5.2 implies that  $\mathfrak{M}_s(G')$  induces a rapidly mixing Markov chain on  $F_r(G)$ . We now introduce a more general class of Markov chains on  $F_r(G)$  and investigate their mixing behaviour. We allow the transition behaviour of the Markov chains to depend on the degree  $d_X(r) =: \|X\|$ ,  $X \in \text{ST}(G')$ , of the newly added node  $r$ , i.e. it may depend on the number of connected components of the forest corresponding to  $X \in \text{ST}(G')$ .

DEFINITION 5.3. Let  $G = (V, E)$  be an undirected graph with extension  $G' = (V', E')$  as in Lemma 5.2. For any  $\lambda \in \mathbb{R}_+$  we define the Markov chain  $\mathfrak{M}_s^\lambda(G') = (X_t)_{t \in \mathbb{N}}$  on  $\text{ST}(G')$  by the transition probabilities: If  $X_t = X \in \text{ST}(G')$  is the state of  $\mathfrak{M}_s(G')$  at time  $t \in \mathbb{N}$ , then we draw uniformly and independently  $e \in X$  and  $f \in E'$  and set

$$Y := (X \setminus \{e\}) \cup \{f\}.$$

If  $Y \in \text{ST}(G')$ , then we set

$$X_{t+1} = \begin{cases} Y & \text{with probability } p, \\ X & \text{with probability } 1 - p, \end{cases}$$

where  $p := \frac{1}{2} \min\{1, \lambda^{\|Y\| - \|X\|}\}$ . If  $Y \notin \text{ST}(G')$ , then we set  $X_{t+1} := X$ . We denote the transition matrix of this chain by  $P_\lambda$ .

In the Markov chain  $\mathfrak{M}_s^\lambda(G')$  for  $0 < \lambda < 1$  transitions to forests with fewer components are preferred, for  $\lambda > 1$  transitions to forests with more components are preferred, while for  $\lambda = 1$  we have the chain of Section 2 on  $G'$ ,  $\mathfrak{M}_s^\lambda(G') = \mathfrak{M}_s(G')$ . A consequence for the stationary distribution is the following.

**THEOREM 5.4.** *For any  $\lambda \in \mathbb{R}^+$  the Markov chain  $\mathfrak{M}_s^\lambda(G')$  is ergodic. The stationary distribution  $\pi_\lambda$  is given by*

$$(5.1) \quad \pi_\lambda(X) := \frac{\lambda^{\|X\|}}{Z(\lambda)}$$

with normalization  $Z(\lambda) := \sum_{X \in \text{ST}(G')} \lambda^{\|X\|}$ .

*Proof.*  $\mathfrak{M}_s^\lambda(G')$  has the same transition graph as  $\mathfrak{M}_s(G')$  and thus is irreducible and aperiodic by the definition of  $\mathfrak{M}_s^\lambda(G')$ . Further for  $X, Y \in \text{ST}(G')$  with  $P_\lambda(X, Y) > 0$ ,

$$\begin{aligned} \pi_\lambda(X)P_\lambda(X, Y) &= \frac{1}{2m(n-1)} \frac{\lambda^{\|X\|}}{Z(\lambda)} \min\{1, \lambda^{\|Y\| - \|X\|}\} \\ &= \frac{1}{2m(n-1)} \frac{1}{Z(\lambda)} \min\{\lambda^{\|X\|}, \lambda^{\|Y\|}\} \\ &= \frac{1}{2m(n-1)} \frac{\lambda^{\|Y\|}}{Z(\lambda)} \min\{\lambda^{\|X\| - \|Y\|}, 1\} \\ &= \pi_\lambda(Y)P_\lambda(Y, X), \end{aligned}$$

with  $n := |V'|$ ,  $m := |E'|$ . Thus  $\pi_\lambda$  is the stationary distribution of  $\mathfrak{M}_s^\lambda(G')$ . ■

In the next theorem we use the canonical paths of Section 4 to estimate the mixing time of  $\mathfrak{M}_s^\lambda(G')$  efficiently. In consequence we have a polynomial sampling scheme for the set of forests of any graph  $G$  with weights proportional to  $\lambda^{d(X)}$ ,  $d(X)$  being the number of connected components of  $X$ .

**THEOREM 5.5.** *Let  $G' = (V', E')$  be the extension of an undirected graph  $G = (V, E)$  and  $\lambda \in \mathbb{R}^+$ . Then the mixing time  $\tau_\lambda$  of the Markov chain  $\mathfrak{M}_s^\lambda(G')$  is bounded by*

$$(5.2) \quad \tau_\lambda(\varepsilon) \leq 2n^2 m \lambda' (n \log(m \lambda') + \log \varepsilon^{-1})$$

for all  $\varepsilon \in (0, 1)$ , where  $\lambda' := \max\{\lambda, \lambda^{-1}\}$ ,  $n := |V'|$  and  $m := |E'|$ .

*Proof.* The Markov chains  $\mathfrak{M}_s^\lambda(G')$  and  $\mathfrak{M}_s(G')$  have the same transition graph. Therefore we can use the canonical paths  $\Gamma_{G'}$  with respect to  $\mathfrak{M}_s(G')$  of Section 4. We have to estimate the congestion measure

$$(5.3) \quad \varrho_\lambda(\Gamma_{G'}) := \max_{\substack{(B,C) \in \text{ST}(G')^2 \\ P_\lambda(B,C) > 0}} \frac{1}{\pi_\lambda(B)P_\lambda(B,C)} \sum_{\gamma_{XY} \in \mathcal{P}(B,C)} \pi_\lambda(X)\pi_\lambda(Y)|\gamma_{XY}|,$$

where  $\mathcal{P}(B, C)$  is the set of canonical paths in  $\Gamma_{G'}$  which use transition  $(B, C)$ . The length of each canonical path is bounded by  $n - 1$ , and  $P_\lambda(B, C) \geq 1/2m(n - 1)\lambda'$  for  $B \neq C$ . Thus from (5.3),

$$\varrho_\lambda(\Gamma_{G'}) \leq \frac{2n^2m\lambda'}{Z(\lambda)} \max_{P_\lambda(B,C) > 0} \sum_{\gamma_{XY} \in \mathcal{P}(B,C)} \frac{\lambda^{\|X\| + \|Y\|}}{\lambda^{\|\bar{B}\|}}.$$

For any fixed transition  $(B, C)$  of  $\mathfrak{M}_s^\lambda(G')$  and  $\gamma_{XY} \in \mathcal{P}(B, C)$ , this transition is coded by  $\bar{B} = X \oplus Y \oplus B$ . As  $B \setminus C$  contains exactly one edge, say  $e \in E'$ , there is by Theorem 4.2 at most one canonical path which codes  $B$  by  $\bar{B}$  and leads from  $B$  directly to  $C$ . Further, by construction of the paths and their codings we have  $B \oplus \bar{B} = X \oplus Y$  as well as  $B \cap \bar{B} = X \cap Y$ . This implies that  $\|B\| + \|\bar{B}\| = \|X\| + \|Y\|$  and we obtain, for all transitions  $(B, C)$  of  $\mathfrak{M}_s^\lambda(G')$ ,

$$\sum_{\gamma_{XY} \in \mathcal{P}(B,C)} \frac{\lambda^{\|X\| + \|Y\|}}{\lambda^{\|\bar{B}\|}} \leq \sum_{\bar{B} \in \text{ST}(G')} \lambda^{\|\bar{B}\|} = Z(\lambda).$$

As a result,

$$(5.4) \quad \varrho_\lambda(\Gamma_{G'}) \leq 2n^2m\lambda',$$

and with  $\hat{\pi}_\lambda := \min_{X \in \text{ST}(G')} \pi_\lambda(X)$ ,

$$(5.5) \quad \tau_\lambda(\varepsilon) \leq 2n^2m\lambda'(\log \hat{\pi}_\lambda^{-1} + \log \varepsilon^{-1}).$$

We have  $\hat{\pi}_\lambda^{-1} \leq \lambda^{-n}Z(\lambda) \leq (\lambda')^n|\text{ST}(G')|$  for  $\lambda < 1$ , and  $\hat{\pi}_\lambda^{-1} \leq Z(\lambda) \leq (\lambda')^n|\text{ST}(G')|$  for  $\lambda \geq 1$ . Since  $|\text{ST}(G')| \leq m^n$ , this implies

$$\tau_\lambda(\varepsilon) \leq 2n^2m\lambda'(n \log(m\lambda') + \log \varepsilon^{-1}). \blacksquare$$

REMARK. For the transition from  $\mathfrak{M}_s^\lambda(G')$  via the bijection  $\text{Sp}$  from Lemma 5.2 to the corresponding Markov chain on  $F_r(G)$  we have to note that  $|V| = |V'| - 1$  and  $|E| = |E'| - |V|$ , which changes the bound only by a polynomial factor. As a consequence we obtain a rapidly mixing Markov chain on the class  $F_r(G)$  of forests with roots or equivalently on the class of spanning trees in a rooted graph  $G'$  with stationary distribution proportional to  $\lambda^{d(X)}$ ,  $d(X)$  being the degree of the root  $r$ . ■

**6. Connected spanning subgraphs.** In this section we consider the class  $S_c(G)$  of connected spanning subgraphs of a graph  $G = (V, E)$ . As any  $X \in S_c(G)$  has node set  $V$  we can identify  $X$  with its node set and

thus identify  $X \subset E$  with subsets of  $E$  which define connected spanning subgraphs.

Also for  $S_c(G)$  no efficient randomized approximation scheme and no rapidly mixing Markov chain with uniform distribution on  $S_c(G)$  as stationary distribution is known (see Welsh and Merino (2000)). We have the following connection of  $S_c(G)$  to spanning trees.

LEMMA 6.1. *For an undirected graph  $G = (V, E)$  there exists a graph  $G'' := (V'', E'')$  such that there exists a bijection between  $\text{ST}(G'')$  and*

$$\bigcup_{A \in S_c(G)} \{(A, T) \mid T \in \text{ST}(A)\}.$$

*Proof.* Let  $V_E := \{v_e \mid e \in E\}$  be a set of new nodes and define  $G'' := (V'', E'')$  with  $V'' := V \cup V_E$  and  $E'' := \bigcup_{e \in E} \{\{v, v_e\} \mid v \in e\}$ , i.e. the edge  $e = \{v, w\}$  is replaced by two edges  $\{v, v_e\}$  and  $\{w, v_e\}$ . For each node  $e \in E$  we denote one end node by  $e_r$  (right end node) and one end node by  $e_l$  (left end node). For  $X \in \text{ST}(G'')$  the subset  $A := \{e \in E \mid \{v_e, e_r\} \in X\}$  is a connected spanning subgraph of  $G$ . Further,  $T := \{e \in E \mid \{v_e, e_r\} \in X \text{ and } \{v_e, e_l\} \in X\}$  is a spanning tree of  $G$ .

Conversely,  $X$  can be reconstructed from  $A$  and  $T$ , since

$$X = \{\{v_e, e_r\} \mid e \in A\} \cup \{\{v_e, e_l\} \mid e \in T \cup (E \setminus A)\}.$$

Thus this mapping is bijective. ■

Thus any  $A \in S_c(G)$  corresponds to as many spanning trees of  $G''$  as  $A$  has itself. We introduce as in Section 5 a weighting on  $\text{ST}(G'')$  by

$$(6.1) \quad \|X\| := |A|$$

where  $X \in \text{ST}(G'')$  corresponds to  $(A, T)$  in Lemma 6.1.

DEFINITION 6.2. For  $\lambda \in \mathbb{R}^+$  we define the Markov chain  $\mathfrak{M}_{S_c}^\lambda(G'') = (X_t)$  on  $\text{ST}(G'')$  as in Definition 5.3 where the norm  $\|X\|$  is defined by (6.1).

Similarly to the argument in Section 5 we obtain (for details see Fehrenbach (2003))

THEOREM 6.3. *The Markov chain  $\mathfrak{M}_{S_c}^\lambda(G'')$  is ergodic for any  $\lambda \in \mathbb{R}^+$  with stationary distribution*

$$(6.2) \quad \pi_\lambda(X) := \frac{\lambda^{\|X\|}}{Z(\lambda)} \quad \text{with} \quad Z(\lambda) := \sum_{X \in \text{ST}(G'')} \lambda^{\|X\|}$$

for  $X \in \text{ST}(G'')$ . The mixing time  $\tau_\lambda$  is bounded by

$$(6.3) \quad \tau_\lambda(\varepsilon) \leq 2n^2 m \lambda' (n \log(m \lambda') + \log \varepsilon^{-1})$$

for all  $\varepsilon \in (0, 1)$  with  $\lambda' := \max\{\lambda, \lambda^{-1}\}$ ,  $n := |V''|$  and  $m := |E''|$ .

Based on Lemma 6.1 we can transfer the Markov chain  $\mathfrak{M}_{S_c}^\lambda(G'')$  to the basic space  $\bigcup_{A \in S_c(G)} \{(A, T) \mid T \in \text{ST}(A)\}$  and project it on  $S_c(G)$ . Then we obtain a Markov chain  $\mathfrak{M}_{S_c}^\lambda(G)$  on  $S_c(G)$ . The weight of a state  $A \in S_c(G)$  with respect to its stationary distribution is given by  $\lambda^{|A|} |\text{ST}(A)|$ . The rapid mixing property persists since the numbers of edges and nodes of  $G$  and  $G''$  only differ by a polynomial factor. Thus we obtain

**COROLLARY 6.4.** *The Markov chain  $\mathfrak{M}_{S_c}^\lambda(G)$  induced by  $\mathfrak{M}_{S_c}^\lambda(G'')$  on  $S_c(G)$  is rapidly mixing with stationary distribution*

$$(6.4) \quad \pi_\lambda(A) = \lambda^{|A|} |\text{ST}(A)|, \quad A \in S_c(G).$$

**REMARK.** As a consequence of Corollary 6.4 we do not get a polynomial randomized approximation scheme for  $|S_c(G)|$  as  $\pi_\lambda$  is not the uniform distribution. We may, however, obtain such a scheme for some functionals like  $\sum_{A \in S_c(G)} |\text{ST}(A)|$ . This can also be obtained directly from  $|\text{ST}(G)|$ , since

$$\sum_{A \in S_c(G)} |\text{ST}(A)| = 2^{m-(n-1)} |\text{ST}(G)|$$

as  $|E - T| = m - (n - 1)$  for all  $T \in \text{ST}(G)$ .

## References

- D. Aldous (1983), *Random walks on finite groups and rapidly mixing Markov chains*, in: Lecture Notes in Math. 986, Springer, 243–297.
- D. Aldous (1990), *The random walk construction of uniform spanning trees and uniform labelled trees*, SIAM J. Discrete Math. 3, 450–465.
- A. Broder (1989), *Generating random spanning trees*, in: Proc. 30th Annual IEEE Symposium on Foundations of Computer Science (FOCS), 442–447.
- R. Cordovil and M. Moreira (1993), *Bases-cobases graphs and polytopes of matroids*, Combinatorica 13, 157–165.
- M. Cryan, M. Dyer, L. A. Goldberg, and M. Jerrum (2002), *Rapidly mixing Markov chains for sampling contingency tables with a constant number of rows*, in: Proc. 43rd Symposium on Foundations of Computer Science (FOCS), 711–720.
- P. Diaconis (1988), *Group Representations in Probability and Statistics*, IMS Lecture Notes Monogr. Ser. 11, Hayward.
- P. Diaconis and D. Stroock (1991), *Geometric bounds for eigenvalues of Markov chains*, Ann. Appl. Probab. 1, 36–61.
- R. Diestel (1996), *Graphentheorie*, Springer, 1996.
- T. Feder and M. Mihail (1992), *Balanced matroids*, in: Proc. 24th Annual ACM Symposium on Theory of Computing, 26–38.
- J. Fehrenbach (2003), *Design und Analyse stochastischer Algorithmen auf kombinatorischen Strukturen*, PhD thesis, Univ. of Freiburg.
- M. Jerrum (1998), *Mathematical foundations of the Markov chain Monte Carlo method*, in: M. Habib et al. (eds.), Probabilistic Methods for Algorithmic Discrete Mathematics, Algorithms Combin. 16, Springer, 116–165.



- M. Jerrum (2003), *Counting, Sampling and Integrating: Algorithms and Complexity*, Birkhäuser.
- M. Jerrum and A. Sinclair (1996), *The Markov chain Monte Carlo method: an approach to approximate counting and integration*, in: D. Hochbaum (ed.), *Approximation Algorithms for NP-hard problems*, PWS Publ., 482–520.
- M. Jerrum and J.-B. Son (2002), *Spectral gap and log-Sobolev constant for balanced matroids*, in: 43rd IEEE Symposium on Foundations of Computer Science, Computer Soc. Press, 721–729.
- M. Jerrum, L. G. Valiant and V. V. Vazirani (1989), *Random generation of combinatorial structures from a uniform distribution*, *Theor. Comput. Sci.* 43, 169–188.
- B. Morris and A. Sinclair (2004), *Random walks on truncated cubes and sampling 0-1 knapsack solutions*, *SIAM J. Comput.* 34, 195–226.
- A. Sinclair (1992), *Improved bounds for mixing rates of Markov chains and multicommodity flow*, *Combin. Probab. Comput.* 1, 351–370.
- A. Sinclair (1993), *Algorithms for Random Generation and Counting: A Markov Chain Approach*, *Progr. Theoret. Comput. Sci.*, Birkhäuser.
- D. Welsh and C. Merino (2000), *The Potts model and the Tutte polynomial*, *Math. Phys.* 41, 1127–1152.

Department of Mathematics  
University of Freiburg  
Eckerstr. 1  
79104 Freiburg, Germany  
E-mail: jfehrenbach@gmx.de  
ruschen@stochastik.uni-freiburg.de

*Received on 3.3.2004;*  
*revised version on 23.8.2005*

(1733)